# Catastrophic Photo-z Errors and the Dark Energy Parameter Estimates with Cosmic Shear

Lei Sun [1,2], Zu-Hui Fan[1], Charling Tao[2], Jean-Paul Kneib[3], Stéphanie Jouvel[3], André Tilquin[2]

sunl@bac.pku.edu.cn

## ABSTRACT

We study the impact of catastrophic errors occurring in the photometric redshifts of galaxies on cosmological parameter estimates with cosmic shear tomography. We consider a fiducial survey with 9-filter set and perform photo-z measurement simulations. It is found that a fraction of 1% galaxies at $z_{\rm spec} \sim 0.4$ is misidentified to be at $z_{\rm phot} \sim 3.5$. We then employ both $\chi^2$ fitting method and the extension of Fisher matrix formalism to evaluate the bias on the equation of state parameters of dark energy, $w_0$-$w_a$, induced by those catastrophic outliers. By comparing the results from both methods, we verify that the estimation of $w_0$-$w_a$ from the fiducial 5-bin tomographic analyses can be significantly biased. We further investigate the requirements of spectroscopic calibration to reduce the bias to the level insignificant compared with statistical errors. For the overall fraction of catastrophic failures $f_{\rm cata} = 1\%$ and the survey area $A = 1000$ deg$^2$, the needed number of spectroscopic redshift measurements for galaxies with their photometric redshifts within the range $z_{\rm phot} = [3, 4]$ is $N_{\rm spec} \geq 350$, and 850, respectively, in order to reduce the joint bias on $w_0$-$w_a$ to be smaller than $2\sigma$ and $1\sigma$, where $\sigma$ represents the joint statistical error of $w_0$-$w_a$. We further give the scaling relation $N_{\rm spec} \propto f_{\rm cata} \times A$. Thus for $f_{\rm cata} = 1\%$ and $A = 10000$ deg$^2$, the $2\sigma$ and $1\sigma$ requirements for $N_{\rm spec}$ are $N_{\rm spec} \geq 3500$ and 8500, respectively.

*Subject headings:* cosmology-gravitational lensing, large scale structure of universe

[1]Department of Astronomy, Peking University, Beijing 100871, China

[2]Centre de Physique des Particules de Marseille, CNRS/IN2P3-Luminy and Université de la Méditerranée, Case 907, F-13288 Marseille Cedex 9, France

[3]Laboratoire d'Astrophysique de Marseille - UMR 6110 - CNRS-Université de Provence-38 rue Fréderic Joliot-Curie, 13013 Marseille, France

## 1.   Introduction

The weak gravitational lensing effect by large scale structures, namely the cosmic shear, is believed to be a powerful tool to measure the cosmological parameters including the equation of state of dark energy. Over the last ten years, weak lensing observations have already provided increasingly better constraints on the matter density $\Omega_M$ and the amplitude of mass fluctuations often represented by $\sigma_8$, the *rms* amplitude of the smoothed and linearly extrapolated mass density perturbations on the scale of 8 Mpc $h^{-1}$ (Hoekstra, Yee & Gladders 2002, Jarvis et al. 2003, Rhodes et al. 2004, Heymans et al. 2004, Hoekstra et al. 2006, Semboloni et al. 2006, Massey et al. 2007, Benjamin et al. 2007, Fu et al. 2008). Their contributions to the constraints on the total neutrino mass have also been investigated (Tereno et al. 2008, Gong et al. 2008, Ichiki et al. 2008, Li et al. 2008). Several more ambitious weak lensing surveys targeting at hundred to thousand times larger areas, such as DES[1], LSST[2], JDEM[3] and EUCLID[4], are being planned aiming at exploring the properties of dark energy.

Future large surveys are expected to be able to control statistical errors to insignificant levels. Therefore different systematics can become dominant sources of contaminations in future weak lensing studies, and should be investigated carefully. These include the effect of the point spread function (PSF) (e.g., Amara & Refregier 2008), the methodology in the shear measurement (Bacon et al.2001, Erben et al. 2001, Hirata & Seljak 2003, Heymans et al. 2006, Massey et al. 2007, Miller et al. 2007, Kitching et al. 2008, Bridle et al. 2008, Ngan et al. 2008, Semboloni et al. 2008), redshift errors (Ma et al. 2006, Huterer et al. 2006), intrinsic alignments of source galaxies (see Heavens 2001 for a review, Bridle & King 2007, Fan 2007), uncertainties in the computation of the non-linear power spectrum (see discussion in Huterer 2002, Van Waerbeke et al. 2002), and so on. In this paper, we focus on the impact of the photometric redshift errors on cosmological studies.

The necessities in measuring vast amount of galaxy redshifts in cosmic shear observations make the multi-color photometric determination of redshifts the only feasible way in obtaining the redshift information of source galaxies. The inaccuracy of photometric redshifts can contaminate considerably the cosmological information embedded in cosmic shears (Huterer et al. 2006; Ma et al. 2006; Amara & Refregier 2007; Wittman et al. 2007; Abdalla

---

[1]http://www.darkenergysurvey.org

[2]http://www.lsst.org

[3]http://jdem.gsfc.nasa.gov

[4]http://sci.esa.int/science-e/www/object/index.cfm?fobjectid=42266

et al. 2008; Margoniner & Wittman 2008; Stabenau et al. 2008). Taking into account the bias and the scatter of $z_{\rm phot}$ with respect to $z_{\rm spec}$, Ma, Hu & Huterer (2006) analyze the corresponding degradation of the constraints on dark energy. It is found that these errors can completely erase the additional information from tomographic lensing observations. In order to fully use of the advantage of the tomographic binning, both the bias and scatter of $z_{\rm phot}$ have to be known to better than 0.003-0.01. Amara & Refregier (2007) and Abdalla et al. (2008) pay special attention to the so-called catastrophic errors in $z_{\rm phot}$, the outliers with $|z_{\rm spec} - z_{\rm phot}| > 1$, and explore their effects on the figure of merit of dark energy parameter constraints.

While most of the existing studies concentrate on the error range of parameter constraints, the systematic bias in the determination of cosmological parameters due to the photometric redshift errors should also be carefully analyzed. A large systematic bias can lead to a completely wrong conclusion on the values of the cosmological parameters although the statistical errors may be small (Amara & Refregier 2008). Our study presented in this paper emphasizes precisely this aspect of effects on dark energy parameter estimates arising from catastrophic failures in photo-z measurements. We focus on a survey with similar design as described in Aldering et al. (2004), with its 9-filter set described also in Jouvel et al. 2008 (in preparation). We analyze the bias on dark energy parameter estimates induced by the catastrophic failures occurred around $(z_{\rm spec}, z_{\rm phot}) \sim (0.4, 3.5)$. We investigate in detail the requirements of the number of spectroscopic redshift measurements to retrieve the correct estimates.

The outline of the paper is as follows. In §2, we introduce the methods to describe the lensing observables and to evaluate the biases. We present the results of bias in §3. In §4, we discuss the calibration requirements to remove the biases due to the catastrophic failures. Discussions and conclusions are given in §5.

## 2. Methodology

In this section, we describe the formalism of lensing observables. Based on simulation results, we then model the source galaxy distribution with catastrophic failures taken into account. Finally we discuss the methods to evaluate the corresponding biases on dark energy parameter estimates, without/with spectroscopic calibration.

## 2.1. Lensing Observables

The convergence power spectrum $P_{ij}^\kappa(\ell)$ for the $i$th and $j$th redshift bin can be expressed as (see Kaiser 1992, Kaiser 1998, Ma et al. 2006)

$$P_{ij}^\kappa(\ell) = \int_0^\infty dz\, W_i(z)\, W_j(z) \frac{H(z)}{D^2(z)}\, P(k_\ell, z)\,, \tag{1}$$

where $H(z)$ is the Hubble parameter and $D(z)$ is the comoving angular diameter distance. The matter power spectrum $P(k_\ell, z)$ with $k_\ell = \ell/D(z)$ is computed from the linear transfer function of Bardeen et al. (1986), and then the non-linear fitting function of Peacock & Dodds (1996). The weighting function $W_i$ is given by

$$
\begin{aligned}
W_i(z) &= \frac{3}{2}\Omega_m \frac{H_0^2 D(z)}{H(z)}(1+z) \\
&\times \int_z^\infty dz'\, n_i(z') \frac{D_{LS}(z, z')}{D(z')}\,,
\end{aligned} \tag{2}
$$

where $D_{LS}(z, z')$ is the angular diameter distance between $z$ and $z'$, and $n_i$ is the galaxy distribution (normalized to unity) of $i$th redshift bin. We employ the overall galaxy distribution of the form (Smail et al. 1994)

$$n(z) \propto z^\alpha \exp\left[-(1.41z/z_{\mathrm{med}})^\beta\right] \tag{3}$$

where $z_{\mathrm{med}}$ is the median redshift of a survey and we take $\alpha = 2$, $\beta = 1.5$.

Considering only the shot noise and the Gaussian sample variance, the covariance matrix of the lensing observables can be written as

$$\mathrm{Cov}\left[C_{ij}^\kappa(\ell'), C_{kl}^\kappa(\ell)\right] = \frac{\delta_{\ell\ell'}}{(2\ell + 1)\, f_{\mathrm{sky}}\, \Delta\ell}\left[C_{ik}^\kappa(\ell)C_{jl}^\kappa(\ell) + C_{il}^\kappa(\ell)C_{jk}^\kappa(\ell)\right]. \tag{4}$$

where $\Delta\ell$ is the band width in multipole $\ell$, and $f_{\mathrm{sky}}$ is the fractional sky coverage of a survey. The total power spectrum $C_{ij}^\kappa$ is given by

$$C_{ij}^\kappa = P_{ij}^\kappa + \delta_{ij}\frac{\sigma_\gamma^2}{n_g^i}\,, \tag{5}$$

where $\sigma_\gamma^2$ is the variance of each component of the intrinsic ellipticity of source galaxies, and $n_g^i$ is the surface number density of galaxies in the $i$th bin.

In our study, we consider a fiducial space-based survey with $f_{\mathrm{sky}} = 0.025$, corresponding to 1000 deg$^2$, the summation of $n_g^i$, i.e., the total surface number density of source galaxies $n_g^{\mathrm{tot}}$, is taken to be 100 arcmin$^{-2}$, $\sigma_\gamma$=0.22 and $z_{\mathrm{med}}$=1.26. We are aware of the likely

extension of the survey area. Thus in analyzing the requirement of the spectroscopic redshift calibration, we derive a scaling relation that allows us to estimate the corresponding calibration requirement for any survey area and any catastrophic error fraction.

Investigations by Ma et al. (2006) show that tomographic analyses with multi-bins of source galaxies (denoted as $N_{\rm bin}$) can improve the statistical errors on cosmological parameter constraints considerably in comparison with those of 2-D analyses. The improvement increases with the increase of $N_{\rm bin}$. On the other hand, we cannot gain much further improvement with $N_{\rm bin} > 5$. Hence throughout the paper, we consider lensing tomography with 5 equally spaced redshift bins within the range of $z = [0, \ 3]$. All galaxies with $z > 3$ are simply added to the last bin. The resulting redshift binning is illustrated in Figure 2.

## 2.2.   Modeling galaxy distribution with catastrophic failures

Photometric estimations of galaxy redshifts through multi-waveband observations largely depend on the characteristic features of the spectral energy distribution (SED) of galaxies, such as the Lyman break and Balmer/4000$\mathring{A}$ break. The locations of the features give us the redshift information of the galaxies. Therefore the accuracy of photometric redshifts is sensitive to the wavelength coverage of observations. It is also affected by the observational signal-to-noise ratio and the estimation method. Besides, the spectroscopic calibration plays an important role in improving the accuracy of photo-z, which shall be discussed specifically in §4.

To realistically assess the accuracy of photometric redshifts for our survey, we first consider a 9-filter set which covers a wavelength range from visible blue to infrared with $\lambda \sim 390$ nm to $\sim 1700$ nm. Then a series of simulations are carried out . The details of the simulations will be presented in Jouvel et al. 2008 (in preparation). Here we give a brief description. The simulations are performed with the 'Le Phare' software. First we generate galaxies according to the GOODS luminosity function by Dahlen et al. (2005), each with an assigned spectrum based on the Coleman Extended templates (Coleman et al. 1980). This set of templates contains 66 spectra obtained by interpolating between 5 standard spectra: elliptical, Sbc, Scd, irregular and star forming galaxies. We then apply the filtering process to the simulated spectra with the 9-filter set. We thus obtain the SED for each simulated galaxy. The magnitudes of galaxies are randomly assigned in accordance with the properties of our fiducial survey. The photometric redshift for each galaxy is then computed using the Le Phare photometric redshift code. The results are shown in Figure 1.

It is seen that scatters and biases around $z_{\rm spec}$ exist for $z_{\rm phot}$. Besides, there are notable

islands in the $z_{\mathrm{spec}}$-$z_{\mathrm{phot}}$ plot. The dominant one is located at $z_{\mathrm{spec}} \sim 0.4$ and $z_{\mathrm{phot}} \sim 3.5$, which accounts for more than 50% of the galaxies with $|z_{\mathrm{phot}} - z_{\mathrm{spec}}| > 1$. This island obviously comes from the confusion of the $4000\mathring{A}$ break with the Lyman break in galaxy SEDs due to the limited wavelength coverage. It is found that the fraction of galaxies with such large catastrophic errors is about 1% of the total number of galaxies. In our study, we particularly analyze the effect of this island on the constraints of dark energy parameters.

To characterize the true distribution of galaxies whose photometric redshifts suffer from catastrophic failures, we use the following Gaussian distribution with the form

$$A_{\mathrm{cata}} \exp\Big[ -\frac{(z - z_m)^2}{2\sigma^2} \Big], \tag{6}$$

where the 'true' mean redshift of the island $z_m = 0.398$ and the standard deviation $\sigma = 0.108$ are obtained from fitting to the spectro(true)-z distribution histogram of the island. The parameter $A_{\mathrm{cata}}$ is determined by the fraction of catastrophic failures $f_{\mathrm{cata}} = 0.01$. Then in a tomographic division according to galaxies' photometric redshifts, this part will be mistakenly shifted from low-z bin (the 'source') to another high-z bin (the 'target'). Specifically, for the $N_{\mathrm{bin}} = 5$ tomography we consider, the 1% galaxies with their true-z falling in the first bin are assigned to the last (5th) bin whose redshift range is $z_{\mathrm{phot}} > 2.4$.

## 2.3.  Evaluation of biases on dark energy parameter estimates

Once the distribution of catastrophic failure fraction is known, we can now evaluate the biases it induces on dark energy parameter estimates.

As mentioned in the previous subsection, we need to be aware of the fact that observationally, source galaxies are assigned to different redshift bins in terms of their 'observed' photometric redshifts rather than their true redshifts. Hence, a fraction of low redshift galaxies can be mistakenly distributed to high-redshift bins due to the catastrophic errors. As a consequence, weak lensing signals from those relevant tomographic bins are perturbed and biases on cosmological parameter estimates may arise, if models used to fit the data were not modified correspondingly.

Concerning the fiducial survey in this study, we have known from simulations that there will be $f_{\mathrm{cata}} = 0.01$ galaxies being misplaced into the 5th bin (denoted as 'Bin-5'), whilst their true redshifts are within the range of the first bin (denoted as 'Bin-1'). The true distribution of this catastrophic failure part can be described with Eq. (6) and will be labeled as Bin-0 for convenience in the following discussion. Note that Bin-0 is within Bin-1.

For Bin-1, when Bin-0 galaxies within it are misplaced in Bin-5 due to catastrophic red-

shift errors, the number of galaxies contained in Bin-1 is reduced. The redshift distribution of the galaxies left in it also changes somewhat in comparison with the true distribution without the catastrophic errors. The number change in Bin-1 does not affect the lensing signal because the galaxy distribution within the bin is always normalized to unity. The shape change of the redshift distribution in the bin may have some effects. On the other hand, however, this change may already be included in the modeled overall redshift distribution of galaxies derived from photometric redshifts. Thus in our consideration, we assume that the misplacement of Bin-0 does not affect the lensing signal from Bin-1 (the 'source' bin). Conversely, as shown in Figure 3, the lensing signal from Bin-5 is perturbed (the 'target' bin). It now contains two populations of galaxies, the original one with the fraction of $1 - \hat{f}$, and the contamination from Bin-0 with the fraction of $\hat{f}$, i.e., $n_5 \rightarrow (1 - \hat{f})n_5 + \hat{f}n_0$. Here, $\hat{f} = f_{\text{cata}}/(f_{\text{cata}} + f_5)$ is the relative fraction of galaxies with catastrophic failures in Bin-5, with $f_5$ being the fraction of galaxies with their true redshifts falling in Bin-5. Note, in Figure 3 we edit out the specific shape information of the redshift distribution of galaxies in Bin-0 and Bin-5 to demonstrate the fractional relation only.

It is now straightforward to evaluate the perturbations on lensing signals, i.e., the convergence power spectra. Besides the auto spetra of Bin-5, those cross spetra with one of the bins is Bin-5 will also get affected. The rest keeps unperturbed. Thus we have (Huterer & Bernstein 2007, private communication)

$$C_{i5}^{\kappa} \rightarrow (1 - \hat{f})C_{i5}^{\kappa} + \hat{f}C_{i0}^{\kappa} \qquad (\text{i} = 1, ..., 4) \tag{7}$$

$$C_{55}^{\kappa} \rightarrow (1 - \hat{f})^2 C_{55}^{\kappa} + 2\hat{f}(1 - \hat{f})C_{05}^{\kappa} + \hat{f}^2 C_{00}^{\kappa} \tag{8}$$

$$C_{ij}^{\kappa} \rightarrow C_{ij}^{\kappa} \qquad (\text{otherwise}, \text{i}, \text{j} = 1, ..., 4). \tag{9}$$

In practical analysis, when calculating the 'theoretically expected' lensing signals for Bin-5, we have no knowledge, in the case without spectroscopic calibration, on the fraction of low redshift galaxies and regard all the galaxies being at high redshifts. Or in the case with inadequate spectroscopic calibration, we have inaccurate knowledge therefore are apt to make unprecise estimation of the fraction. This is clearly shown in Figure 3. Consequently systematic biases arise when the theoretically expected signals are compared with the observed ones to constrain cosmological parameters.

As an approximation, a simple extension of the Fisher matrix formalism can be used to compute the bias in cosmological parameters with (Huterer et al. 2006; Amara & Refregier 2008)

$$\delta p_i = F_{ij}^{-1} \sum_{\ell,\alpha,\beta} \left[ C_\alpha^\kappa(\ell) - \bar{C}_\alpha^\kappa(\ell) \right] \mathrm{Cov}^{-1} \left[ C_\alpha^\kappa(\ell), C_\beta^\kappa(\ell) \right] \frac{\partial \bar{C}_\beta^\kappa(\ell)}{\partial p_j} \tag{10}$$

where $\alpha$ and $\beta$ denote a pair of redshift bins. The term $\left[ C_\alpha^\kappa(\ell) - \bar{C}_\alpha^\kappa(\ell) \right]$ reflects the bias on the shear signals, and are calculated from Eqs. (7)-(9) as (rhs-lhs) in the case without calibration[5]. For the cases with calibrations, the lhs terms should be modified to be the same form as rhs terms but with various calibrated $\hat{f}$ , and then the subtraction (rhs-lhs) is performed on these new terms. The covariance of the cross power spectra is given in Eq. (4). $F_{ij}^{-1}$ is the inverse of the Fisher matrix

$$F_{ij} = \sum_\ell \left( \frac{\partial C}{\partial p_i} \right)^T \mathrm{Cov}^{-1} \frac{\partial C}{\partial p_j}, \tag{11}$$

where $C$ is the column matrix of the convergence power spectra and $\mathrm{Cov}^{-1}$ is the inverse of the covariance matrix.

Eq. (10) is a linear approximation around the fiducial model. On the other hand, the true extent of bias can be quantified by a direct chi-square fitting method, by minimizing

$$\chi^2 = \sum_{\ell,\alpha,\beta} \left[ C_\alpha^\kappa(\ell) - \bar{C}_\alpha^\kappa(\ell) \right] \mathrm{Cov}^{-1} \left[ C_\alpha^\kappa(\ell), C_\beta^\kappa(\ell) \right] \left[ C_\beta^\kappa(\ell) - \bar{C}_\beta^\kappa(\ell) \right]. \tag{12}$$

In our analyses, we will utilize both bias evaluation methods. The comparison of their results permits us to quantify the extent to which the linear approximation is valid. We shall come back to this in the following section, associated with the specific study for our fiducial survey.

Throughout our study, a 7 parameter fiducial flat cosmological model is assumed. For equation of state of dark energy, we consider the form $w(z) = w_0 + w_a z/(1+z)$ (Linder 2003). The fiducial values of the 7 parameters are: $\Omega_m = 0.27, w_0 = -1, w_a = 0, h = 0.72, \sigma_8 = 0.78, \Omega_b = 0.0446, n = 0.96$, respectively.

---

[5]Note, here we assume a smooth distribution $n_5(z_{\mathrm{phot}})$ for the calculation of lhs, namely, the model predicted signals. This is in accordance with what people generally do in practice. While in fact there may be a notable bump around $z_{\mathrm{phot}} = 3.5$ (correlated with $n_5(z_{\mathrm{phot}})$) on histogram of $n(z_{\mathrm{phot}})$, due to $z_{\mathrm{phot}}$ being blindly used and Bin-0 being misidentified at high-z, people are not apt to catch the small bump-feature when fitting the histogram to give $n(z_{\mathrm{phot}})$ curve. Our analyses show that this omission of the high-z bump in modeling $n(z_{\mathrm{phot}})$ actually helps to diminish somewhat the bias due to catastrophic failures.

All the weak lensing related calculations are done with the ICOSMO package (Refregier, version 2005[6]), with modifications from 2D weak lensing to tomographic form for our specific study. The $\chi^2$ fitting tool is MPFIT[7] from C. Markwardt.

## 3. The bias on dark energy parameter estimates

To demonstrate clearly their significance, in this section, we present the results of bias on dark energy parameters due to catastrophic failures if there is no spectroscopic calibration for the problematic Bin-5. We apply both methods by $\chi^2$ fitting and by computation with the extension of Fisher matrix formalism. In the next section, we will focus on the requirements for the spectroscopic calibrations.

First, we employ the $\chi^2$ fitting method. As described in the previous section, the observed lensing signals have the form given by the right-hand side of Eqs. (7)-(9), whilst the model-predicted signals used to fit the data are computed with the left-hand side terms. Particularly for Bin-5, the observed signal is a combination of the two parts illustrated in the top-left panel of Figure 3. On the other hand, the model calculation is done with the galaxy distribution illustrated in the bottom-left panel of Figure 3. The $\chi^2$ fitting is then applied to find the 'best-fit' values of the cosmological parameters by comparing the model-predicted lensing signals with the 'observational data'. In this fitting process, the 'measurement errors' for each 'data point' are theoretically calculated using a binned form of Eq. (4).

In our calculations, the $\ell$ range is taken to be $[50, 3000]$. We discard weak lensing information from all those scales beyond $\ell = 3000$ to avoid complications from various small-scale effects, such as baryonic cooling (White 2004, Zhan & Knox 2005, Jing et al. 2006) and non-Gaussianity (White & Hu 2000, Cooray & Hu 2001). We apply binning procedures to $\ell$. According to Hu & White 2001, the bin width should be at least twice as big as $\ell_{\text{field}}$, in order to reduce the covariance between the bins arising from the survey geometry. We then take 10 bins being uniform in logarithmic scale over the range $\ell = [50, 3000]$.

In the right panel of Figure 4, we show the resulting bias on the equation of state of dark energy obtained from $\chi^2$ fitting. Gaussian priors $\sigma(\Omega_b) = 0.01$ for $\Omega_b$ and $\sigma(p_i) = 0.05$ for others are applied upon the hidden parameters in our 7-parameter fiducial model. These priors are consistent with the current constraints. It is found that the 1% catastrophic failure fraction can bring large biases on $w_0$ and $w_a$ in the tomographic 5 z-bins case, highly

---

[6]for a latest version, see http://www.icosmo.org/ and Refregier et al. (2008)

[7]http://cow.physics.wisc.edu/ craigm/idl/fitting.html

significant compared to their statistical errors. The 'best fit' values are $w_0 = -3.5$ and $w_a = 3.8$ in comparison with the fiducial ones $w_0 = -1$ and $w_a = 0$.

The results from the extended Fisher matrix formalism are shown in the right panel of Figure 4. The same $\ell$-binning process and Gaussian priors are applied on. The 'best fit' values are found to be $w_0 = -4.2$ and $w_a = 6$. Thus the linear approximation results notably larger bias than those of the $\chi^2$ fitting. This significant discrepancy reveals clearly the necessities in going beyond the linear approximation from the Fisher matrix formalism.

In summary, our analyses show that for our fiducial survey with a 9-filter set, the catastrophic photo-z errors with $f_{\rm cata} \sim 1\%$ can bias the dark energy parameters far exceeding the statistical errors if no further spectroscopic calibrations are conducted.

Our calculations are done for the survey area $A = 1000 \text{ deg}^2$, or the sky coverage $f_{\rm sky} = 0.025$. Increasing the survey area reduces the statistical errors with $\sigma \propto 1/\sqrt{f_{\rm sky}}$. On the other hand, the absolute biases on cosmological parameters should not change much. This can be seen clearly from Eq. (10) in which $f_{\rm sky}$ enters through $F_{ij}^{-1}$ and $\text{Cov}^{-1}$. Since both are proportional to $f_{\rm sky}^{-1}$, their dependencies on $f_{\rm sky}$ cancel out. Thus with a larger sky coverage, the effect of the bias relative to the statistical errors would be even larger if the catastrophic fraction $f_{\rm cata}$ remains to be $\sim 1\%$.

## 4.    Requirements of Spectroscopic Calibrations

The large bias seen in Figure 4 clearly reveals the necessities to take extra measures to reduce the effects of the catastrophic redshift errors. Among others, the spectroscopic calibration can play crucial roles in this regard. It is known that the spectroscopic calibration can be done either by taking spectra of a sub-sample of source galaxies after photo-z measurements or by employing a sophisticated technique where the spectral information is included in the determination of photometric redshifts. We consider the former case in our analyses, and explore the requirements of the spectroscopic calibration. Specifically, we study the needed number of galaxies with measured spectroscopic redshifts in order to retrieve the unbiased parameter estimates.

Since we particularly analyze the 'island' catastrophic failures around $(z_{\rm spec}, z_{\rm phot}) \sim (0.4, 3.5)$ seen in Figure 1, we consider the calibration for galaxies with their photo-z in the range $3 < z_{\rm phot} < 4$. We assume that our simulation results shown in Figure 1 reflect the real relation between photo-z and the true redshifts for observed galaxies. In other words, we take the simulation results as the fiducial distribution with an overall $f_{\rm cata} = 1\%$. We then investigate how many spectra are needed for the sub-sample of galaxies with $3 < z_{\rm phot} < 4$ in

order to gain the information on the catastrophic contamination, mainly the value of $f_{\mathrm{cata}}$.

Our analyzing procedures are as follows. We use the simulated galaxies with $3 \leq z_{\mathrm{phot}} \leq 4$ as our parent fiducial sample. Then for a given $N_{\mathrm{spec}}$, where $N_{\mathrm{spec}}$ is the number of galaxies chosen for spectral measurements, we perform a large number of Monte Carlo random samplings of $N_{\mathrm{spec}}$ galaxies from the parent catalog. For each realization, we first compute the estimated fraction of catastrophic failures $\bar{f}_{\mathrm{cata}}$ by counting the number of galaxies among $N_{\mathrm{spec}}$ that are identified in the realization to be in Bin-0, taking into account the overall fraction of galaxies that have $3 \leq z_{\mathrm{phot}} \leq 4$ . With this sub-sample of Bin-0 galaxies, we then fit their distribution to the form given by Eq. (6) to obtain the estimated $z_m$ and $\sigma$, denoted as $\bar{z}_m$ and $\bar{\sigma}$. With these calibrated values of $\bar{f}_{\mathrm{cata}}$, $\bar{z}_m$ and $\bar{\sigma}$, we are able to calculate the model-predicted signals. By comparing them with those from the fiducial distribution, we therefore can evaluate the biases on cosmological parameters for each individual realization. The biases arise from the differences between the calibrated $(\bar{f}_{\mathrm{cata}}, \bar{z}_m, \bar{\sigma})$ and the fiducial $(f_{\mathrm{cata}}, z_m, \sigma)$. With the results from all the realizations, we can further evaluate statistically the level of bias from $N_{\mathrm{spec}}$ calibrations.

Figure 5 shows the results on $w_0$-$w_a$ of 100 realizations for $N_{\mathrm{spec}} = 100$ (upper panels) and $N_{\mathrm{spec}} = 500$ (lower panels), respectively. The symbols are the 'best fit' values from individual realizations, and the statistical errors are calculated at the fiducial values for the purpose of clarity. The left panels are for the results of full calculations including both the auto- and cross- correlations between different redshift bins. The right panels are the results considering only the auto-correlations within individual redshift bins. In each panel, the pluses and triangles correspond, respectively, to the results from $\chi^2$ fitting and from the Fisher matrix calculations of Eq. (10). First when the bias is large and the fitted values are outside the $3\sigma$ contour, the results from the linear approximation of Eq. (10) start to deviate from the $\chi^2$ fitting results, as seen most evidently in the upper right panel. This shows the limitation of the linear approximation. If in some cases, very large biases are expected, one is necessary to perform a full $\chi^2$ fitting to evaluate quantitatively the biases. Otherwise, an over-estimate of the bias on $w_0$-$w_a$ would be obtained from the linear analysis. This is in good accordance with the results shown in Figure 4. On the other hand, when the biases are within the $3\sigma$ statistical contour, the two sets of results are in excellent agreement. Thus in this regime, the linear approximation of Eq. (10) provides us an accurate and convenient way to estimate the bias effects. We further find that the triangles from the linear calculations are almost perfectly on a straight line with little scatters. The results from the full $\chi^2$ calculations follow the same line when the biases are relatively small, and are bent downward when the biases are large. Noting that the biases depend on three quantities $(\bar{f}_{\mathrm{cata}}, \bar{z}_m, \bar{\sigma})$, the lining up of the results indicate that there is a dominant factor among the three. Our detailed analyses, presented later in Figure 6, show that the biases are mainly

determined by $(\bar{f}_{\text{cata}} - f_{\text{cata}})$, and the shape of the redshift distribution in Bin-0 (described by the parameters $\bar{z}_m$, $\bar{\sigma}$) has insignificant effects.

The straight line behavior presents us a quantitative way to describe the joint bias of $w_0$-$w_a$. We define the distance between the fiducial point and the best-fit point of a realization as the $w_0$-$w_a$ joint bias, denoted by $\delta(w_0\text{-}w_a)$, for this individual realization. Furthermore, to quantify the relative bias with respect to the statistical errors, we define the distance between the fiducial point and the intersection point of the straight line with a statistical error contour as the corresponding $w_0$-$w_a$ joint statistical error. This definition accounts correctly for the correlation between $w_0$ and $w_a$.

To see the dependence of the bias on calibrations, in Figure 6, we show the correlation of the joint bias $\delta(w_0\text{-}w_a)$ defined above and the residual $f_{\text{cata}}$ (corresponding to $f_{\text{cata}} - \bar{f}_{\text{cata}}$) from each individual realization for $N_{\text{spec}} = 500$. It clearly demonstrates that the bias scales approximately linearly with the fraction of catastrophic failures. The small dispersions of $\delta(w_0\text{-}w_a)$ at a fixed residual $f_{\text{cata}}$ correspond to results with different calibrated $\bar{z}_m$ and $\bar{\sigma}$ for Bin-0. Thus this reveals quantitatively that the effect of the specific distribution within Bin-0 is sub-dominate, in comparison with the impact of overall fraction of Bin-0, namely $\bar{f}_{\text{cata}}$. This fact actually allows us to obtain a simple scaling relation for the dependence of the bias on the survey area $f_{\text{sky}}$ and on the catastrophic fraction $f_{\text{cata}}$, which will be discussed shortly.

With the biases obtained from individual realizations, we can statistically quantify the typical level of bias for a given $N_{\text{spec}}$. From Figure 5 and 6, we see that the average bias from all the realizations of a given $N_{\text{spec}}$ is expected to be close to zero. Hence we use the corresponding root-mean-square bias calculated from the bias distribution to characterize the bias level for that particular $N_{\text{spec}}$. In Figure 7, we show the dependence of the rms bias $\sigma[\delta(w_0\text{-}w_a)]$ on $N_{\text{spec}}$. For each $N_{\text{spec}}$, we conduct 500 realizations. The bias for each individual realization is calculated by employing the Fisher matrix formalism Eq. (10) since it can give good approximation when the residual $f_{\text{cata}}$ is reasonably small. We can see that $\sigma[\delta(w_0\text{-}w_a)] \propto 1/\sqrt{N_{\text{spec}}}$. About 850 and 350 spectroscopic measurements for galaxies with $z_{\text{phot}} \sim [3, 4]$ are needed so that the rms bias $\sigma[\delta(w_0\text{-}w_a)]$ can be smaller than $1\sigma$ and $2\sigma$ statistical errors, respectively.

The specific results on the required $N_{\text{spec}}$ shown above are calculated with $f_{\text{sky}} = 0.025$ and $f_{\text{cata}} = 1\%$. However, based on our analyses, we can derive a simple scaling relation which permits us to estimate the needed $N_{\text{spec}}$ for any $f_{\text{sky}}$ and $f_{\text{cata}}$. As illustrated in Figure 6, the bias is linearly proportional to $\bar{f}_{\text{cata}} - f_{\text{cata}}$. Thus the rms bias $\sigma[\delta(w_0\text{-}w_a)]$ should be approximately proportional to the rms of $\bar{f}_{\text{cata}} - f_{\text{cata}}$, namely $\sigma(\bar{f}_{\text{cata}} - f_{\text{cata}})$. Furthermore, for a set of realizations, the number of calibrating galaxies that are identified in Bin-0 follows

the Poisson distribution with a mean value of $f_{\rm cata} \times N_{\rm tot}/\Delta N_5 \times N_{\rm spec}$, which is also the value of the variance, where $N_{\rm tot}$ and $\Delta N_5$ represent the total number of galaxies and the number of galaxies with $3 \leq z_{\rm phot} \leq 4$. Then it is clear that $\sigma(\bar{f}_{\rm cata} - f_{\rm cata}) \propto \sqrt{f_{\rm cata}}/\sqrt{N_{\rm spec}}$. On the other hand, as we discussed in §3, the absolute bias should not depend on the survey area. The statistical error $\sigma(w_0\text{-}w_a)$ scales as $1/\sqrt{f_{\rm sky}}$. Therefore the relative bias with respect to the statistical error scales as $\sqrt{f_{\rm cata} \times f_{\rm sky}/N_{\rm spec}}$. This in turn gives rise to the scaling relation for the needed $N_{\rm spec} \propto f_{\rm cata} \times f_{\rm sky}$. We then have

$$N_{\rm spec} > 850 \left( \frac{f_{\rm cata}}{0.01} \right) \left( \frac{f_{\rm sky}}{0.025} \right) \tag{13}$$

for keeping the bias smaller than $1\sigma$ statistical error. Thus for $f_{\rm cata} = 1\%$ and $f_{\rm sky} = 0.25$ ($A = 10,000$ deg$^2$), we would need $N_{\rm spec} > 8500$.

## 5. Discussion

In this paper, we concentrate on the island part of the galaxies in the $(z_{\rm spec}, z_{\rm phot})$ plot where $z_{\rm phot}$ is completely mis-estimated. This type of large catastrophic failures in the photometric redshift measurements arise from the confusion of different characteristics of galaxies' SED due to the limited wavelength coverage. The most notable features used in the photo-z measurement of galaxies are the Lyman and $4000\mathring{A}$ breaks. The mis-identification of the two results the islands at $(z_{\rm spec}, z_{\rm phot}) \sim (0.4, 3.5)$ and $(z_{\rm spec}, z_{\rm phot}) \sim (3, 0.2)$ seen in Figure 1. We particularly analyze the systematic bias induced by such islands on cosmological parameter estimations. We find that a fraction of $f_{\rm cata} = 1\%$ catastrophic failures due to galaxies at $z_{\rm spec} \sim 0.4$ being misidentified to be at $z_{\rm phot} \sim 3.5$, which is the dominate case of catastrophic failure occurrence in our photo-z simulations, can bring systematic biases on $w_0$ and $w_a$ at the level far exceeding the statistical errors in 5-tomographic-bin analysis.

It has been realized that the spectroscopic redshift calibration may be necessary for future weak lensing surveys in order to reduce the uncertainty of photometric redshift errors. The large bias shown in our analyses reinforces such an argument. We investigate the requirements of the spectroscopic calibration to retrieve the unbiased parameter estimates. We find that for $f_{\rm cata} = 1\%$, and $f_{\rm sky} = 0.025$, about 400-1000 spectral redshift measurements for galaxies with $z_{\rm phot} = [3, 4]$ are needed. We further present a scaling relation of $N_{\rm spec} \propto f_{\rm cata} \times f_{\rm sky}$. Thus for a future survey with a survey area of $10,000$ deg$^2$, the size of the calibration sample in the redshift range $z_{\rm phot} = [3, 4]$ should be on the order of $10^4$ if $f_{\rm cata}$ remains to be $\sim 1\%$. In analyzing the error degradation from general photometric errors, people have argued that the inaccuracy of the high redshift photometric measurements can be compensated by increasing the measurement accuracy for galaxies at relatively low redshifts.

Therefore much less spectroscopic calibration samples at high redshifts are needed (e.g., Ma et al. 2006). Concerning the catastrophic-error-induced bias discussed in this paper, however, there is no such a compensation. Thus the required $\sim 10^4$ spectral calibrations in $z_{\rm phot} = [3,\ 4]$ may be regarded as challenging tasks.

It is also argued that the inclusion of both U and near-IR filters is essential for the accuracy of photo-z and the minimization of the catastrophic outliers (e.g. Dahlen et al. 2008, Abdalla et al. 2008), as the Lyman and $4000\mathring{A}$ breaks can be effectively distinguished. Given that our simulations with 9-filter set already include the near-IR bands, we consider the improvements from U band coverage. Extending the wavelength coverage to $3200\mathring{A}$, our simulations show that the catastrophic failure fraction decreases dramatically to the level around 0.1% (Jouvel et al. 2008, in preparation). In this case, the bias should no longer be significant compared to the statistical errors. At the same time, it may not be straightforward to integrate U band for space-based telescopes.

It has also been proposed to filter out those problematic galaxies, such as late type galaxies, low signal-to-noise galaxies, or very low/high redshift (e.g., $z < 0.5/z > 2.5$) galaxies whose photo-zs are most suspicious, in the shear analysis to reduce the impact of catastrophic failures (Abdalla et al. 2008; Margoniner & Wittman 2008; Heavens, Kitching & Taylor 2006). In such cases, the statistical power of a survey is inevitably lowered to certain extents depending on the fraction of galaxies discarded. In Figure 8, we show the degradation of the constraints if those galaxies at $z < 0.5$ and $z > 2.5$ (thus 3 bins left) are excluded in the analyses. We see that the Figure of Merit (denoted as F.o.M)[8] decreases from 41 to 24, i.e., almost a factor of two degradation. If we re-divide the galaxies with $0.5 \leq z \leq 2.5$ into 5 finner bins (right panel of Figure 8), the F.o.M = 29, or a degradation factor $\sim 1.5$.

In our current analyses, we isolate the island galaxies appeared around $(z_{\rm spec}, z_{\rm phot}) = (0.4, 3.5)$ to see clearly their effects on cosmic shear analysis. On the other hand, the scatters and small bias of $z_{\rm phot}$ around $z_{\rm spec}$ have been modeled and studied in, for example, Ma et al. (2006). With $\sigma_z$ and $z_{\rm bias}$ as well as their uncertainties $\Delta\sigma_z$ and $\Delta z_{\rm bias}$ taken into account, the statistical errors of the constrained cosmological parameters will be degraded depending on the values of $(\sigma_z, z_{\rm bias}, \Delta\sigma_z, \Delta z_{\rm bias})$. For instance, considering $(\sigma_z, z_{\rm bias}, \Delta\sigma_z, \Delta z_{\rm bias}) = (0.05(1+z), 0, 0.003, 0.003)$, the degradation factor for $w_0$ and $w_a$ is about 1.5 (Ma et al. 2006). Then a crude estimate based on our analyses gives the needed $N_{\rm spec}$ at $z \sim [3,\ 4]$ to be about $8500/(1.5)^2 \sim 4000$ to control the bias down to $1\sigma$ level for $f_{\rm cata} = 1\%$ and

---

[8]F.o.M $= \sqrt{\det(\mathrm{F})}$, where F is the reduced Fisher matrix for $w_0$ and $w_a$ after marginalization over other parameters.

$f_{\rm sky} = 0.25$. More complete analyses should be done for both the degradation and the bias simultaneously taking into account both the small and the large photo-z errors.

## REFERENCES

Abdalla, F. B., Amara, A., Capak, P., Cypriano, E. S., Lahav, O., & Rhodes, J. 2008, MNRAS, 387, 969

Aldering, G. et al. 2004, PASP, submitted (astro-ph/0405232)

Amara, A. & Refregier, A. 2007, MNRAS, 381, 1018

Amara, A. & Refregier, A. 2008, MNRAS, 391, 228

Bacon, D. J., Refregier, A., Clowe, D., & Ellis, R. 2001, MNRAS, 325, 1065

Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15

Bridle, S. & King, L. 2007, New Journal of Physics, Volume 9, Issue 12, pp. 444 (astro-ph/0705.0166)

Bridle, S. et al. 2008, astro-ph/0802.1214

Benjamin, J., Heymans, C., Semboloni, E., van Waerbeke, L., Hoekstra, H., Erben, T., Gladders, M. D., Hetterscheidt, M., Mellier, Y., & Yee, H. K. C. 2007, MNRAS, 381, 702

Coleman, G. D., Wu, C.-C., & Weedman, D. W. 1980, ApJS, 43, 393

Cooray, A. & Hu, W. 2001, ApJ, 554, 56

Dahlen, T., Mobasher, B., Somerville, R. S., Moustakas, L. A., Dickinson, M., Ferguson, H. C., & Giavalisco, M. 2005, ApJ, 631, 126

Dahlen, T., Mobasher, B., Jouvel, S., Kneib, J. P., Ilbert, O., Arnouts, S., Bernstein, G., & Rhodes, J. 2008, AJ, 136, 1361

Erben, T., Van Waerbeke, L., Bertin, E., Mellier, Y., & Schneider, P. 2001, A&A, 366, 717

Fan, Z.-H. 2007, ApJ, 669, 10

Fu, L. et al. 2008, A&A, 479, 9

Gong, Y., Zhang, T.-J., Lan, T., & Chen, X.-L. 2008, astro-ph/0810.3572

Heavens, A. F. 2001, Intrinsic Galaxy Alignments and Weak Gravitational Lensing. Yale Worksh. Shapes Galaxies Haloes, May. astro-ph/0109063

Heavens, A. F., Kitching, T. D., & Taylor, A. N. 2006, MNRAS, 373, 105

Heymans, C., Brown, M., Heavens, A., Meisenheimer, K., Taylor, A & Wolf, C. 2004, MNRAS, 347, 895

Heymans, C. et al. 2006, MNRAS, 368, 1323

Hirata, C. & Seljak, U. 2003, MNRAS, 343, 459

Hoekstra, H., Yee, H. K. C., & Gladders, M. D. 2002, ApJ, 577, 595

Hoekstra, H., Mellier, Y., van Waerbeke, L., Semboloni, E., Fu, L., Hudson, M. J., Parker, L. C., Tereno, I., & Benabed, K. 2006, ApJ, 647, 116

Hu, W. & White, M. 2001, ApJ, 554, 67

Huterer, D. 2002, Phys. Rev. D, 65, 3001

Huterer, D. & Bernstein, G. 2007, private communication

Huterer, D., Takada, M., Bernstein, G., & Jain, B. 2006, MNRAS, 366, 101

Ichiki, K., Takada, M., & Takahashi, T. 2008, astro-ph/0810.4921

Jarvis, M., Bernstein, G., Jain, B., Fischer, P., Smith, D., Tyson, J. A. & Wittman, D. 2003, AJ, 125, 1014

Jing, Y.-P., Zhang, P.-J., Lin, W.-P., Gao, L., & Springel, V. 2006, ApJ, 640, 119

Kaiser, N. 1992, ApJ, 388, 272

Kaiser, N. 1998, ApJ, 498, 26

Kitching, T. D., Miller, L., Heymans, C. E., van Waerbeke, L., Heavens, A. F. 2008, MNRAS, 390, 149

Li, H., Liu, J., Xia, J.-Q., Sun, L., Fan, Z.-H., Tao, C., Tilquin, A., & Zhang, X.-M. 2008, astro-ph/0812.1672

Linder E. V., 2003, Phys. Rev. Lett., 90, 091301

Ma, Z., Hu, W., & Huterer, D. 2006, ApJ, 636, 21

Margoniner, V. E. & Wittman, D. M. 2008, ApJ, 649, 31

Massey, R. et al. 2007, ApJS, 172, 239

Massey, R. et al. 2007, MNRAS, 376, 13

Miller, L., Kitching, T. D., Heymans, C., Heavens, A. F., & van Waerbeke, L. 2007, MNRAS, 382, 315

Ngan, W.-H. W., Van Waerbeke, L., Mahdavi, A., Heymans, C., & Hoekstra, H. 2008, astro-ph/0809.3465

Peacock, J. A., & Dodds, S. J. 1996, MNRAS, 280, 19

Refregier, A., Amara, A., Kitching, T., & Rassat, A. 2008, astro-ph/0810.1285

Rhodes, J., Refregier, A., Collins, N. R., Gardner, J. P., Groth, E. J., & Hill, R. S. 2004, ApJ, 605, 29

Semboloni, E., Mellier, Y., van Waerbeke, L., Hoekstra, H., Tereno, I., Benabed, K., Gwyn, S. D. J., Fu, L., Hudson, M. J., Maoli, R., & Parker, L. C. 2006, A&A, 452, 51

Semboloni, E., Tereno, I., Van Waerbeke, L., & Heymans, C. 2008, astro-ph/0812.1881

Smail, I., Ellis, R. S., & Fitchett, M. J. 1994, MNRAS, 270, 245

Stabenau, H. F., Connolly, A., & Jain, B. 2008, MNRAS, 387, 1215

Tereno, I., Schimd, C., Uzan, U. Z., Kilbinger, M., Vincent, F. H., & Fu, L. 2008, astro-ph/0810.0555

van Waerbeke, L., Mellier, Y., Pello, R., Pen, U.-L., McCracken, H. J., & Jain, B. 2002, A&A, 393, 369

Wittman, D. M., Riechers, P., & Margoniner, V. E. 2007, ApJ, 671, 109

White, M. 2004, Astropart. Phys., 22, 211

White, M. & Hu, W. 2000, ApJ, 537, 1
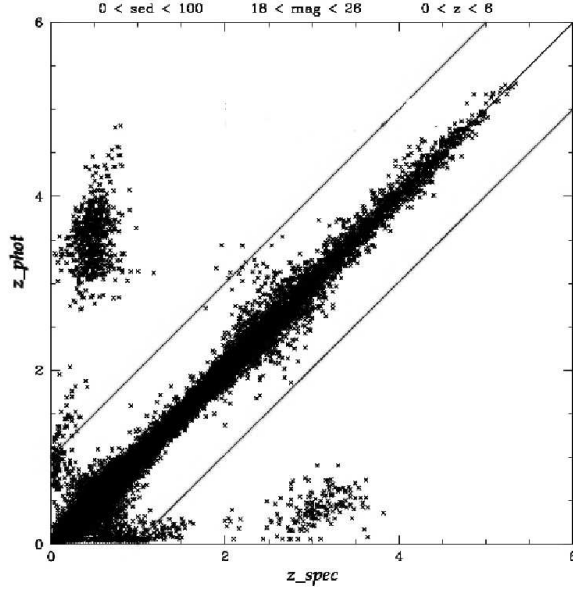
Zhan, H. & Knox, L. 2005, ApJ, 616, 75

---

Fig. 1.— The result of Le Phare simulation for our fiducial survey with a 9-filter set. A magnitude cut of 26 AB is set to fit the median redshift of the survey. The 2 grey lines show the borders inside which the relation $|z_{\mathrm{spec}} - z_{\mathrm{phot}}| < 1$ holds.
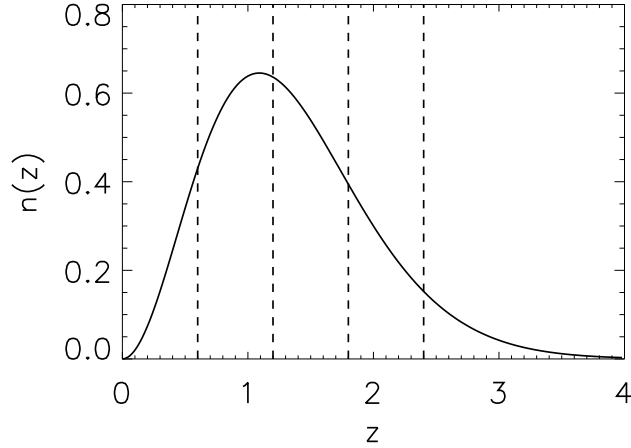


Fig. 2.— Source galaxy distribution $n(z)$. The solid curve is the overall galaxy distribution defined in Equation (3). 5 tomogr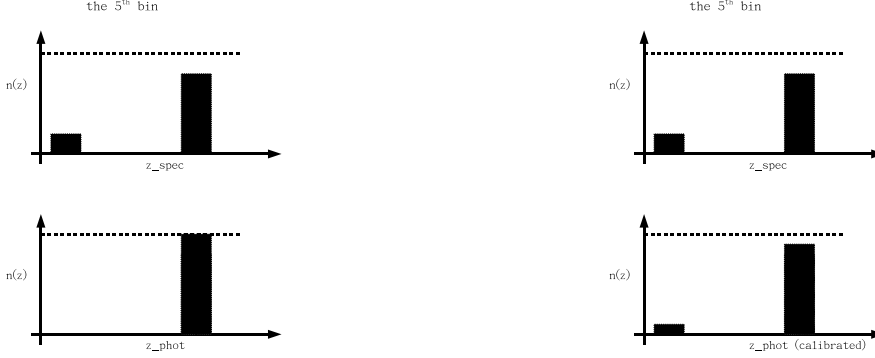aphic bins are equally spaced between $0 < z < 3$, denoted by dashed vertical lines. All galaxies with $z > 3$ are simply added into the last bin.

Fig. 3.— A sketch to explain where the bias comes from. Left panels and right panels shows the cases without and with calibration, respectively. The true distribution of Bin-5 contains 2 populations as shown in the top 2 identical panels. Bottom-left: the 'observed' distribution of Bin-5 without calibration information. Bottom-right: an example of the calibrated distribution of Bin-5, derived from an individual realization of calibration sampling. Here the specific shapes of distributions are edited out to demonstrate the fractional relation only.
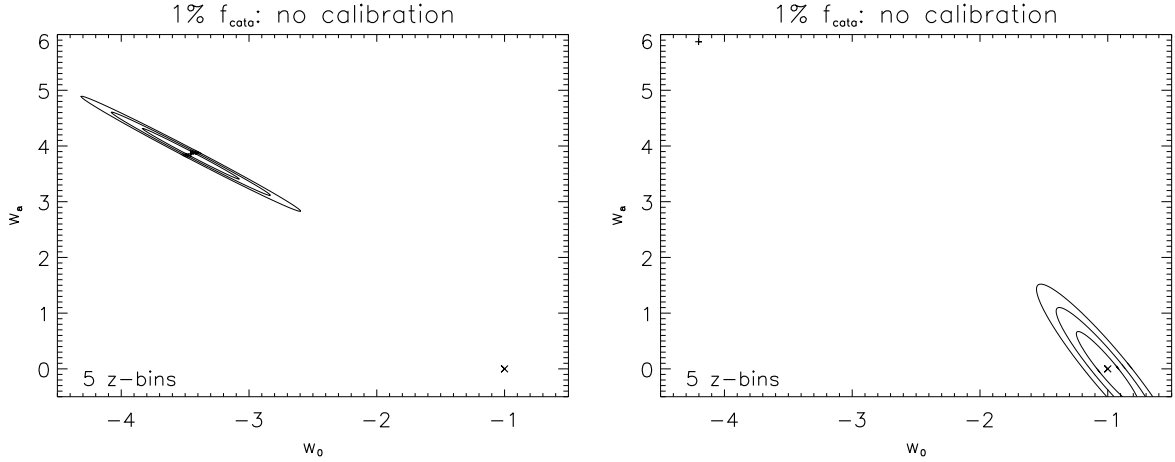


Fig. 4.— Biases on dark energy parameters $w_0$ and $w_a$ in 5 z-bins tomography, from $\chi^2$ fitting (left panel) and from Fisher matrix calculations (right panel), respectively. The fiducial value is denoted by the oblique cross, and the plus represents the 'best fit' value. The error contours $(1, 2, 3\sigma$, inside-out) are given at the best fit value for fitting result and at the fiducial value for Fisher matrix calculations, respectively. Gaussian priors $\sigma(\Omega_b) = 0.01$ for $\Omega_b$ and $\sigma(p_i) = 0.05$ are applied upon all other hidden parameters.
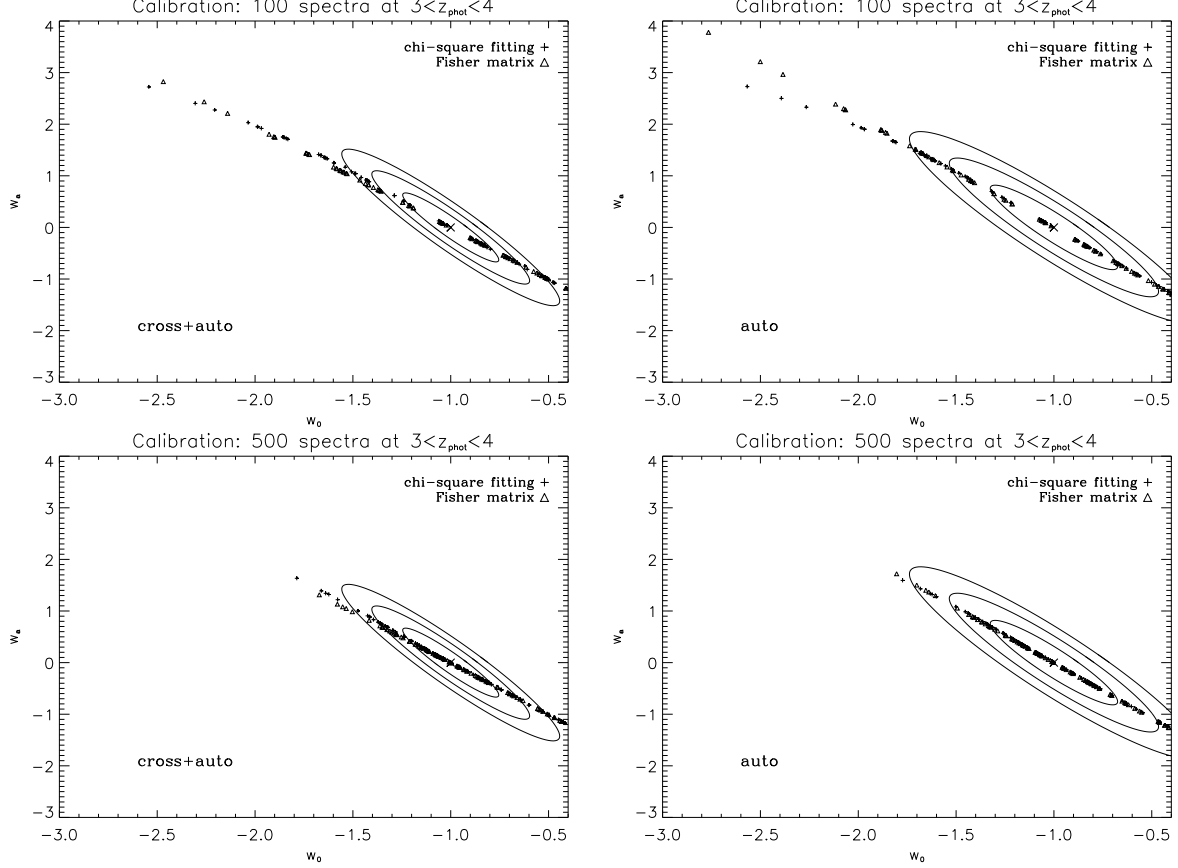
Fig. 5.— Biases on $w_0$-$w_a$ from 100 realizations for $N_{\rm spec} = 100$ (upper panels) and $N_{\rm spec} = 500$ (lower panels), respectively. Left panels: the results of full calculations including both the auto- and cross- correlations between different redshift bins. Right panels: the results considering only the auto-correlations within individual redshift bins. In each panel, the pluses and triangles denote the 'best fit' values of individual realizations from $\chi^2$ fitting and from the Fisher matrix calculations, respectively. The contours correspond to $1, 2, 3\sigma$ statistical errors computed at the fiducial values.
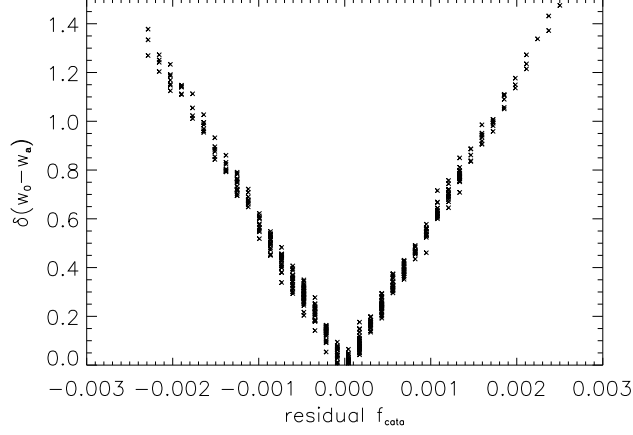
Fig. 6.— Joint $w_0$-$w_a$ bias, $\delta(w_0$-$w_a)$, and the residual $f_{\text{cata}}$ (corresponding to $f_{\text{cata}} - \bar{f}_{\text{cata}}$) from different realizations, for $N_{\text{spec}} = 500$. The dispersion of $\delta(w_0$-$w_a)$ at a fixed residual $f_{\text{cata}}$ reflects the effect of different $\bar{z}_m$ and $\bar{\sigma}$ estimated from individual realizations.
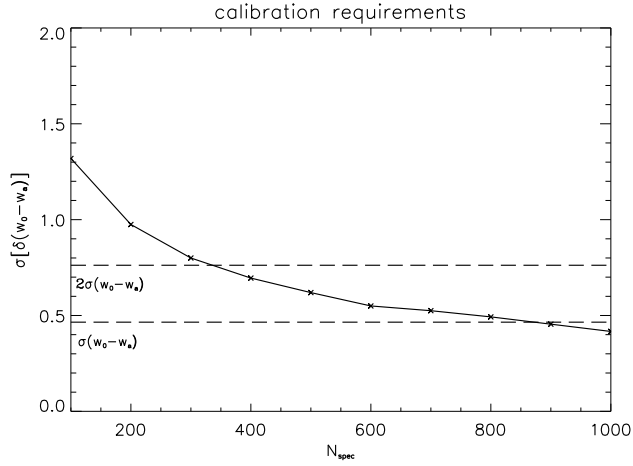


Fig. 7.— Relation of the rms bias, $\sigma[\delta(w_0$-$w_a)]$, with the number of spectra at $3 \leq z_{\text{phot}} \leq 4$ available for calibration, $N_{\text{spec}}$. The dashed lines denote the joint $1\sigma$ and $2\sigma$ statistical errors of $w_0$-$w_a$, respectively.
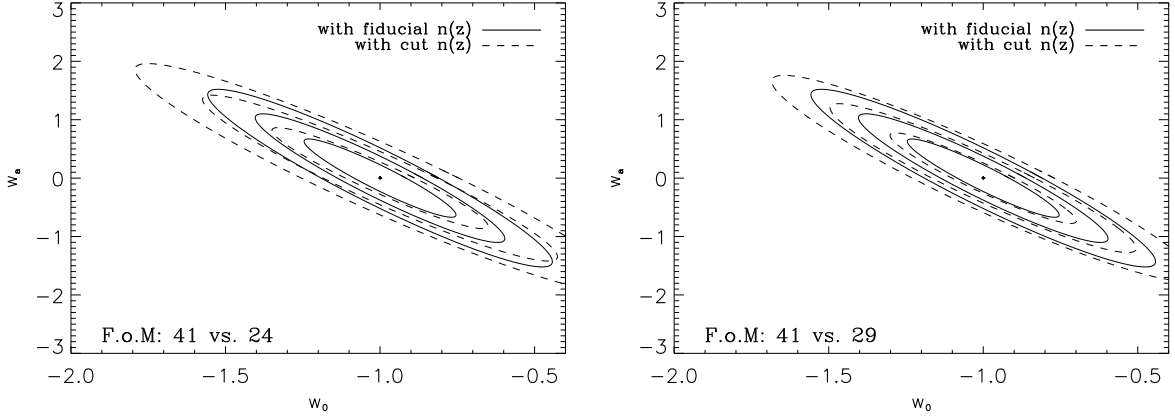
Fig. 8.— The comparisons of constraints on $w_0$-$w_a$. The solid contours (in both panels): the results for the fiducial 5 bin tomography. The dashed contours: the results for discarding those galaxies at $z < 0.5$ and $z > 2.5$, either with the left 3 bins at z=[0.5, 2.5] kept unchanged (in left panel) or with galaxies at z=[0.5, 2.5] re-devided into 5 finner bins (in right panel). Gaussian priors $\sigma(\Omega_b) = 0.01$ for $\Omega_b$ and $\sigma(p_i) = 0.05$ are applied upon all other hidden parameters.